IN THE CLAIMS:

The text of all pending claims, (including withdrawn claims) is set forth below. Cancelled and not entered claims are indicated with claim number and status only. The claims as listed below show added text with underlining and deleted text with ~~strikethrough~~. The status of each claim is indicated with one of (original), (currently amended), (cancelled), (withdrawn), (new), (previously presented), or (not entered).

Please AMEND claim 24 in accordance with the following:

Claims 1-21 (canceled)

22. (previously presented) A method for computer-aided determination of a sequence of actions for a system having states, the method comprising the steps of:

performing a transition in state between two states on the basis of an action;

determining the sequence of actions to be performed such that a sequence of states results from the sequence of actions;

optimizing the sequence of steps with regard to a prescribed optimization function, including a variable parameter; and

using the variable parameter to set a risk which the resulting sequence of states has with respect to a prescribed state of the system.

23. (previously presented) The method as claimed in claim 22, further comprising the step of:

using approximative dynamic programming for the purpose of determination.

24. (currently amended) The method as claimed in claim 23, further comprising the step of:

basing the approximative dynamic programming on Q-learning.

25. (previously presented) The method as claimed in claim 24, further comprising the steps of:

forming an optimization function with Q-learning in accordance with the following rule:

$OFQ = Q(x; w^a)$, and

adapting weights of the function approximator in accordance with the following rule:

$$w_{t+1}^{a_t} = w_t^{a_t} + \eta_t \cdot \aleph^{\kappa}(d_t) \cdot \nabla Q(x_t; w_t^{a_t})$$

wherein

$$d_t = r(x_t, a_t, x_{t+1}) + \gamma \max_{a \in A} Q(x_{t+1}, w_t^{a_t}) - Q(x_t, w_t^{a_t})$$

26. (previously presented) The method as claimed in claim 23, further comprising the step of:

basing the approximative dynamic programming on TD(λ)-learning.

27. (previously presented) The method as claimed in claim 26, further comprising the steps of:

forming the optimization function within TD(λ)-learning in accordance with the following rule:

OFTD = J(x; w); and

adapting weights of the function approximator are adapted in accordance with the following rule:

$w_{t+1} = w_t + \eta_t \cdot \aleph^{\kappa}(d_t) \cdot z_t$, wherein

$d_t = r(w_t, a_t, x_{t+1}) + \gamma J(x_{t+1}; w_t) - J(x_t; w_t)$, $z_t = \lambda \cdot \gamma \cdot z_{t-1} + \nabla J(x_t; w_t)$, and

$z_{-1} = 0$

28. (previously presented) The method as claimed in claim 27, further comprising the step of:

using a technical system to determine the sequence of actions before the determination measured values are measured.

29. (previously presented) The method as claimed in claim 28, further comprising the step of:

subjecting the technical system to open-loop control in accordance with the sequence of actions.

30. (previously presented) The method as claimed in claim 28, further comprising the step of:

subjecting the technical system to closed-loop control in accordance with the sequence of actions.

31. (previously presented) The method as claimed in claim 30, further comprising the step of:

modeling the system as a Markov Decision Problem.

32. (previously presented) The method as claimed in claim 31, further comprising the step of:

using the system in a traffic management system.

33. (previously presented) The method as claimed in claim 31, further comprising the step of:

using the system in a communications system.

34. (previously presented) The method as claimed in claim 31, further comprising the step of:

using the system to carry out access control in a communications network.

35. (previously presented) The method as claimed in claim 31, further comprising the step of:

using the system to carry out routing in a communications network.

36. (previously presented) A system for determining a sequence of actions for a system having states, wherein a transition in state between two states is performed on the basis of an action, the system comprising:

a processor for determining a sequence of actions, whereby a sequence of states resulting from the sequence of actions is optimized with regard to a prescribed optimization function, and the optimization function includes a variable parameter for setting a risk which the resulting sequence of states has with respect to a prescribed state of the system.

37. (previously presented) The system as claimed in claim 36, wherein the processor is used to subject a technical system to open-loop control.

38. (previously presented) The system as claimed in claim 36, wherein the processor Is used to subject a technical system to closed-loop control.

39. (previously presented) The system as claimed in claim 36, wherein the processor is used in a traffic management system.

40. (previously presented) The system as claimed in claim 36, wherein the processor is used in a communication system.

41. (previously presented) The system as claimed in claim 36, wherein the processor is used to carry out access control in a communications network.

42. (previously presented) The system as claimed in claim 36, wherein the processor is used to carry out routing in a communications network.